

```

1 ggaaggagca ggaggcagcc caggcggagc gggaggagct ggcagcgggg cgcattgccag
61 gcgggggtcc tcagggcgcc ccagccgcgc cgggcggcgg cggcgtgagc caccgcgcag
121 gcagccggga ttgcttacca cctgcagcgt gctttcggag gcggcggctg gcacggaggc
181 cgggctacat gagaagctcg acagggcctg ggatcgggtt cctttcccca gcagtgggca
241 cactgttccg gttcccagga ggggtgtctg gcgaggagtc ccaccactcg gagtccaggg
301 ccagacagtg tggccttgac tcgagaggcc tcttgggtcc gagccctgtt tccaagagtg
361 cagcagcccc tactgtgacc tctgtgagag gaacctcggc gcaactttggg attcagctca
421 gaggtggcac cagattgcct gacaggctta gctggccgtg tgcccttggg agtgctgggt
481 ggcagcaaga gtttgacagc atggatagtt ctgagaccct ggacgccagc tgggaggcag
541 cctgcagcga tggagcaagg cgtgtccggg cagcaggctc tctgccatca gcagagttga
601 gtagcaacag ctgcagccct ggctgtggcc ctgagggtcc cccaacccct cctggctctc
661 acagtgcctt tacctcaagc tttagcttta ttccggtctc gcttgggtct gccggggaac
721 gtggagaagc agaaggctgc ccaccatcca gagaggctga gtccatttgc cagagccccc
781 aggagatggg agccaaagct gccagcttgg acgggcctca cgaggaccgc cgatgtctct
841 ctcagccctt cagtctcttg gctacacggg tctctgcaga cttggcccag gccgcaagga
901 acagctccag gccagagcgt gacatgcatt ctttaccaga catggaccct ggctcctcca
961 gttctctgga tccctcactg gctggctgtg gtggtgatgg gagcagcggc tcaggggatg
1021 cccactcttg ggacaccctg ctcaggaaat gggagccagt gctgcgggac tgctgtctga
1081 gaaaccggag gcagatggag gtaatatcct taagattaaa acttcagaaa cttcaggaag
1141 atgcagttga gaatgatgat tatgataaag ctgagacggt acaacaaaga ttagaagacc
1201 tggaacaaga gaaaatcagc ctgcactttc aacttccttc aaggcagcca gctcttagca
1261 gtttctctgg tcacctggca gcacaagtcc aggtctgcct gcgccttggg gccactcagc
1321 aggccagcgg agatgacacc cacacccccc tgagaatgga gccgaggtcg ttggaaccca
1381 ctgctcagga cagcttgcac gtgtccatca cgagacgaga ctggtctctt caggaaaagc
1441 agcagctaca gaaagaaatc gaagctctcc aagcaaggat gtttgtgtcg gaagccaaag
1501 atcaacagct gagaagggaa atagaggagc aagagcagca actccagtgg cagggtctcg
1561 acctgacccc actggtgggc cagctgtccc tgggtcagct gcaggaggtc agcaaggcct
1621 tgcaggacac cctggcctca gccggtcaga ttcccttcca tgcagagcca ccggaaacca
1681 taaggagcct ccaggaaaaga ataaaaatccc tcaacttgtc acttaaagaa atcactacta
1741 aggtgtgtat gagtgagaaa ttctgcagca ccctgaggaa gaaagttaac gatattgaaa
1801 cccaactacc agccttgctt gaagccaaaa tgcattgcat atcaggaaac catttctgga
1861 cggctaaaga cctcaccgag gagattagat cattaacatc agagagagaa gggctggagg
1921 gactcctcag caagctgttg gtgttgagtt ccaggaatgt caaaaagctg ggaagtgtta
1981 aagaagatta caacagactg agaagagaag tggagcacca ggagactgcc tatgaaacaa
2041 gtgtgaagga aaatactatg aagtacatgg aaacacttaa gaataaactg tgcagctgca
2101 agtgtccact gcttgggaaa gtgtgggaag ctgacttggg agcttgtcga ttgcttatcc
2161 agtgcttaca gctccaggaa gccaggggaa gcctgtctgt agaagatgag aggcagatgg
2221 atgacttaga gggagctgct cctcctattc ccccaggct ccaactccag gataaaagga
2281 agaccctttt gaaggtattg gaagaatgga agactcacct catccctctc ctgactgtg
2341 ctggaggtga acagaaagag gaatcttaca tcctttctgc agaacttggg gaaaagtgtg
2401 aagacatagg caagaagcta ttgtacttgg aagatcaact tcacacagca atccacagtc
2461 atgatgaaga tctcattcag tctctcagga gggagctcca gatggtgaag gaaactctgc

```

FIG. 1A

2521 aggccatgat cctgcagctc cagccagcaa aggaggcggg agaaagagaa gctgcagctt
 2581 cctgcatgac agctgggtgc cacgaagcac aagcctgagg agtgacggga tgggggaggg
 2641 aggtggggcca ccatgtttgg acccgggggg ctgctcttcc ctccccgcgc atagctaaga
 2701 tgcctgaatc aattacggag atacagagcc ttgaggctct tcagtggaaa ggtgggtcat
 2761 gttcattctc atcagtgtga aactgaggag tctgcaattt ggaatatgga gagagagact
 2821 gatttgctga atttccttct aaatgtcact caaaaatttc ttttccatgt cattcttggg
 2881 aatgtcttcc acaggatttg agaatagttt catctcagcc cccattagag agaagttggg
 2941 gtgaattctg gaaaaatgtc tctttttcct gtgccatttg ccttctgctg caacgaaaat
 3001 atttctgat tcaagattct ataaaaagga aaccaagcat aagactctgt catcatacct
 3061 gttacacgtt cctacaggtg cacaatctaa gagagctaata taacctcaga gtctggagtt
 3121 aacagctttt caccttactt ctctgtgat ctaataattat cttagaaaaa ttaatatgca
 3181 atttccaaaa gatattttgg taagacaaca acctcccagt gatatgccac ctttcaattt
 3241 tccttttgg gcaatgattg catctgaaga aaggatccct gagagtctct gtttcatcag
 3301 gacattctga aatttaccga cagtgaggct gtggatggat caggggacct gtataaaatg
 3361 tttgagcctg ttccattttc cctggaacc tgtttcactc aatgccaggc agtgcagcat
 3421 ttaggaaagc agtgcagtac tcagtaaggc agtgcagtac tcagtaacac aatacagtac
 3481 tcaggcagtg cagtactcag taagacagtg cagtgtcag taaggcagtg caatactcag
 3541 taacagtgtg gtactcagta acagtgaagt actcagtaat acagtacagt attcagtaag
 3601 gcagtgaagt actcagtaat acaatacagt actcagttag gcagtgcagt actcaggaat
 3661 gcagcacagt actcaggcag tgcaatactc agtgcggtac tcagtaacac agtgcagtac
 3721 tcagtaacag tgcagtactc agtaacagtg cagtactcag taaggcagtg cagtactcag
 3781 taacacagtg cagtactcag taatacagta cagtactcag taaggcagta tgggtactcag
 3841 taaagcaatg caatgctcag taacacagtg cagtgtcaa taaggcagtg cagtgtcag
 3901 taaggcagtg aattgcttag taacacagtg tagtgtcag taggacagca tagtactcag
 3961 taacacaggg cagctagtac tcagtactat aagtactgag tacttatata ggcaatgtag
 4021 tactcagtaa atcagtgcag tactcagtaa tgcaagggca tttcaggctc ctgctgggct
 4081 gcttctttgg ccagctggg actcctattg agacagctgc aaaacaggct gatttcaatt
 4141 aggcagcact tcccaaagtg cactgaggaa ggtggcccca agagaagctc tctaaacaaa
 4201 ggagtaccct ctctgggtcaa gtacctttgg taaatacacc ataccataat atctgcttgg
 4261 agaaccacaa tgcacattag catattagtc tgagagagaa cttatagtaa ggaaactcac
 4321 ttgattttat ctaacctcaa actttccaag tttaatggat cgtgaatttt tttcatgtaa
 4381 ctctatttca tatcccatag atctagtatt gtacagcact gcattctctg aggaagtccc
 4441 agtccaaact ctgatttaca tcactttaga aaccacactc acacttttgc agagtgttga
 4501 gcttaataac tacctgccac agattggtaa atttaatcca gtggttgttc tgtttgtgct
 4561 tctgtttctca tttatgtgtt tagggatagt gaggttcctg ccttccactag gatccacgga
 4621 tatgagacca tttttgtcat ttctgaagt cacactggcg tttccagaag gcatctgggtg
 4681 ctttgtcag ccttccatgc tgtgcagcac ttctgtcctc agtcaaggag atggccatgc
 4741 ttaagccagc aattggctgg ggtccaggaa acaaagcaaa agcacaatat gtgaatgtgc

FIG. 1B

4801 tgattgtggt ccctatggct ttatctcgag caaaatacac tctacatatt ttaataataa
 4861 gtataattag cttgttcctg gacttcattt tcaatgatga accaaattcc tgaattatth
 4921 ataattgtgt ctaaagaaaa ttatgaactg gtcacatggc acttggaaac cttgagttaa
 4981 ttccagtga gcaaaacttg ggaagagtca ggattggcca cattgccaat aacaaattcc
 5041 tacttcgaca tatgtctttt caaaaagcct cccagacaca agacatctta accgtcacta
 5101 gcccagtgt tttgtattac tcagacacca tcatgaaata attctgtgag gtcatgatgt
 5161 atttgaaaat tctgcaagtt aataactgcc ttgaattgtt tgaacccgaa ataagggttc
 5221 tttggtacct ctagtagata gtgtgttcat ttccctgctg caaatthtga agtatttggg
 5281 caggtgagtc atgtthtaac cacaagccat aactcatctg ttgtctthgc ttggtcttag
 5341 agtatcattc agaaagtccg ctaagggccg gcgtgcttct tctggctaca caaccttctc
 5401 aggacaagcc cactgtctta agccactttg accctgggag acacaggact gtgtatcttc
 5461 aatcatacta tacagcagtt tttgtcaggg gaacataaaa atatccaaga gaggttaggg
 5521 cttagattta aaagcatcaa aacaacaaca atggaaaatt atgttggcga tagccaagac
 5581 cacaagcaaa agcacatact ggaaatgatg agttagaatc tgatttgact gggatgttht
 5641 atgagaatgt aagtgtgata ttatactgtc tgccttgctg gaatgctggc thtcaaagtg
 5701 tcacccattt ttctthcact ggctgagtt aggacatgct atcagtaata gtcccagttc
 5761 catccaactt tctgaaattt cattththt tttgagatgg agtctctctc tgtcacccaa
 5821 gttggagtgc agtggccccg caatctcggc tcaactgcaac ctctacctcc caggttcaag
 5881 ctattctact gctcagcct cccaagtagc tggggttaca ggcatthgac accgggacct
 5941 gatgatttht gtattthtag tagagacagg gcttcaccat gttggctagg aggtctcaa
 6001 actcctgacc tcaggcgatc caccaccacc tcggcttccc aaagtgtgga gattacaggt
 6061 gtgagccacc gcaccgggcc aactthctga aattthcaaaa ctgaattgat cctthctcaa
 6121 attagtatat actattggaa acttgcttht ccctgcagta aggttggtth cccacccca
 6181 gaaacatgta acggttggtg ccatgctaag cccttgccat gctaagccct ttacagtcac
 6241 atcctataat ccccatatca acctataag gaaggtgtth gtagatgatg caactgagcc
 6301 ttaagaggac taattccctt thtctaaggc acagagctgg taaaatgtga agtaatagt
 6361 aacctaacag tcagagacag gcagcatgct cttaactagt gctctthcta agttctctth
 6421 aatgtcttht tgagattthg agccatggaa ctthactgtt cacctggcta agaactcatg
 6481 gccactgtgg aaactctggg tagggagtca aagaaactga gcctggggca aacgaggctt
 6541 cccactgtgc caggggagcc tcaactgtga gtctaggctc agacaggcat caacaaacct
 6601 attcacccca ccatcatcct gatctaacca thccccagtc atcccaggaa aaccactcac
 6661 agcctgacac tgggctgact thcttgaaga thctcatcca attggtgtth thcagaagt
 6721 thccaatatt atgaattctg tgttggtggag aaaagcaacc atgcattthc tgggtcaatgc
 6781 cthctgtgat atgtaattca atactththc ththtaatac thcacctthc ctaacttht
 6841 aattthgtca tgaatttht tgcttcatta aggtthctth ttgttatata aaataaaagc
 6901 tgatatccaa ggcaaaaaaa aaaaaaaaaa

FIG. 1C

1 mppgggpggap aaaggggvsh ragsrdclpp aacfrrrrrla rrpgymsst gpgigflspa
61 vgtlfrfpgg vsgeeshhse srarqcglds rgllvrspvs ksaaaptvts vrgtsahfgi
121 qlrggtrlpd rlswpcgpgs agwqgefaam dssetldasw eaacsdgarr vraagslpsa
181 elssnscspg cgpevpptpp ghsaftssf sfirlslgsa gergeaegcp psreaeshcq
241 spqemgaaaa sldgphedpr clsqpfslla trvsadlaqa arnssrperd mhsldpmdpg
301 sssslpsla gcggdgssgs gdahswdtll rkweplrdc llrnrrqmev islrklqlkl
361 gedavenddy dkaetlqgrl edlegekisl hfqlpsrqp lssflghlaa qvqaalrrga
421 tqgasgddth tplrmeprll eptaqdsilhv sitrrdwlq ekqqlqkeie algarmfvle
481 akdgqlrrei eegeqqlgwq gcdltplvgq lslgqlqevs kalqdtlasa gqipfhaepp
541 etirslqeri kslnlslkei ttkvcmsekf cstlrkkvnd ietqlpalle akmhaisgnh
601 fwtakdltee irsltsereq legllskllv lssrnvkklg svkedyrnrlr revehgetay
661 etsvkentmk ymetlknklc sckcpilgkv weadleacr1 liqclqlqea rgslsveder
721 qmddlegaap pipprlhse krktplkvle ewkthlipsl hcaggeqkee syilsaelge
781 kcedigkkl1 yledqlhtai hshdedliqs lrrelqmvke tlqamilqlq pakeagerea
841 aascmtagvh eaqa

FIG. 2

1 ttgtggaaag gatggatggg tgggtggatg gatggatgga tggatggctg gatggatgga
 61 tagaaatgca acttcaagat ctggttctac agtattacat atgaagcaac tgagactcag
 121 ataaaattca gtgagttgct gaagactatt cactcattga gcctcacatt ctgatgctat
 181 atatatcagg tgctatgcag atgattaggg aagggcagga ctgagattaa actgacaccc
 241 cctacatcct aatccaggct ttttccactg tatggcgatg cctcccaaaa gaattatcac
 301 agcgacttaa taaaggaggg agaaagttag ttgcaaattt caagcttaga aattcacagg
 361 caaaaatggc tgaagggaaa taatcacact aatagtgatg caaggatgta actggaaagg
 421 aaaaaatgga gagaaattaa aaacacaaac agatttagtc agggctctga taaagtgata
 481 atgtaatcaa taaggcaaat caaagataaa tgattagaga tttaaaactt cagcctatca
 541 gaaaaacctg agggaaaaga ttttgaattt tttttatttc ttccaaaaga tcttatttga
 601 catatttttt taaatctaaa tgaccaaaaga tttcgaattt ggaaaggcag acacatatct
 661 atttcccata taaacatcta catttcttgc cacatagcac cgcaacactc atctggttcc
 721 aaagcagacg gcttggccct ctgttatcaa gagcactatc accgttctgg cccggagagc
 781 taagttttct tatacatagg tcgtagccaa agctcagagg aaagaagtga catttggtga
 841 gtgtacagga tatgttctcc agccactttt taggtgctaa gttacatctt gaaataatgt
 901 cagaaattgg aagatgccgt atttaaactt cactagccca agccctaatt cacctcgaca
 961 gtttaatatc ttaatgtttc ttaaaatttg aggactctta gaatgacccc tttaaaaatc
 1021 tgctgaaaaa ctggatcttt ttgtgggaaa aactgtccat acctacatat ctaaaatgtg
 1081 tatgtgggtt cagtctctcc tgtgcacagt cctcctgct ttgaagggat cccaacatg
 1141 tttttttcct agaggtagtt ggtcaccaacc atcggataac atgcttatca cctgggtgat
 1201 aaaatattct ggacaccaa ccccgatgac atgcaattta cctatataaa gaacctgcac
 1261 atgtattcct gaacctaaaa taaaaagtgt taaaaaataa taatacaagt aaaagagctt
 1321 tctgcatgtg ttataaatct gttcaggagg actttccact atgctttcaa gagaagtgtc
 1381 ttgaagaatt tagcatcact ctgaattccc aaaaagctct gtgcttttca aggtgctttt
 1441 gaggaagaag gatgatgctt caactgtctt ctttgctgtg acttctagga agttcattat
 1501 cacatcatat tctcaaccaa ataattaaca cttgacattg gcattttctt ctgtctttgt
 1561 ttttaaatct tattgttgtt tgctaatttt attgttttta tctttgttgt atgtattgtc
 1621 ccggtctctc atttcctccc tctcctgcc aataaaagag acccatagat tttattcttt
 1681 tgttctactt tttcatcaat gtttcttcta catatatctg acatatatgg aaaaaagaat
 1741 tactgcctat atatgcagaa cttgtaactc acaggatttc taaactaaaa gacttttctt
 1801 aataggtgtg aagtgaatt ctgcaagtag ctgcatctgg gaacttcaga gttaaaccga
 1861 tatctcagta taatctgtca ctgctatagt ttcttttctt tgacctccaa acctatgtt
 1921 gaaatttgac cccaatgtt gaaggtagag cctaattggga ggtgtctggg tctgggggca
 1981 gatccctcaa gaatagcttg gtcatcaatg agtgagttct cactctaata gtttctggga
 2041 gagctggttg ttaaaaagag tgtgtcatgt ccttctctc tctttttgt tctctctca
 2101 ccagtgtatc gcctaggcca gctcccctc atcttctgcc atgagtggaa gcagcctgag
 2161 gccctcacca gaagcagatg ttggcactat gcttcttaga cagcctgcaa aactgtgagc
 2221 caaagaattc ctttttcttt ataaatgacc cagcctcagg tattctctta tagcaacgca
 2281 gatggactaa gacactcatg cttcctttca ttagctatct caagctcaaa aattgtgaga
 2341 ctgtgccaaag aattagggca ggatcttgtc attgagaact tacaactttt ctatctaaag

FIG. 3A

2401 caaaaggtcc taatcccaaa ccataattgg aaaggtgcct ctactccaac ct cattgccc
 2461 ttcaactgcc ctttctctgc ctccattcc cctcttactc accgaaagac aaatgagttt
 2521 aggctgggca cagttagcca tcgttctcag aaaaacaaaa tcttgctgct tggattctc
 2581 acttagaatc cttcactta gcttatggat caccctacac agattcctga cagctcactg
 2641 ggccatggaa ccagccagat ttttgagcta catctgcagt atcttttcta ct caggcagc
 2701 tggcataagg tgggaaaatt cattaacatg ctaatgatgt cctattttaa attctagtcc
 2761 tgattattct gtgattaaat ctctcccctt gacgtcgggt tcctttatag ctgagttaca
 2821 gtgcgaaatc aacaatatta ataaagtaca atacagatta ccaatttata cacaagggat
 2881 ggtgaaaaac tgctgcgtaa atctgagggc tgccttttta acttgaaaag cagttgtagg
 2941 ccgggcgcgg tggctcacgc ttgtaattcc agcactttgg gaggcctaag caggtggatc
 3001 acctgaggtc aggagtttga gaccagcctg gccaacatgg tgaaaccccg tcaactacta
 3061 aaaaatataa aaattagctg gcctataatc ccagctacct gggaggctga ggcaggagaa
 3121 tcgcttgaac ctgggagacg gaggttgacg tgagcagagt tcgcgccatt acacttcagg
 3181 ctggttgcaa gaagcaaaac ttttgtctca aaaagagaat gtaaaagcag atgtagcaaa
 3241 cttgtaaaag ttgctgattt ttttcattcc gcccttcttt aaataaacac acacattgga
 3301 atccacaata attaaatggt tatagttaat ttccataaat attcaaggct tggagatag
 3361 atttagcttt tcattctatt tcaactgact taacattgct ccgaattcaa atgtgttacc
 3421 ctgaagtctc aagaacagct gacaggagtt aaagcacact tgtttatcat ctctttatag
 3481 agattgtgtc ttggcaacaa aagccaatgg attcgggtgat gaggtctgtc tccctttgcc
 3541 cttgagtacc ttggcatatt tcagttcctt ttctatcacc gtattgttct attactgaca
 3601 ctccaaactg aatatccaac tcagacaact caagcacctt ccaagaaaga caaacttacc
 3661 actaccctca ttaatttgga aaatgaataa agaagcagaa gattcaaggc ccagatctgc
 3721 aatctaaaaa gcaggaagag aggactcatt catgtgtcat ttaacatct gggcccagta
 3781 aatgaacaaa tattgatggt atagtatatt tactatttta aaagacatca tgctggacac
 3841 caaagcagta cagtgaacc taagaacaat gttctgacct cggtggactc acaatctagc
 3901 cagagacaca agatacaatg atattttaag tgatgcatgc agcaaattca tctcaggctg
 3961 atcattagag atacacttca catcattaaa aggtctctgt acgagagaca ctgctaattg
 4021 ttccctaaga ctgcaaatct ccccttttgc tatggaaaca gaattcatca ctttagtggg
 4081 gcacattgct gccagggtta aggacagctc ccagcccctc ttgcaagtag ctgttaccat
 4141 ataacaaggc taaggctaca gggatgtggc cgtagcagag atgtgtgcag atttctaaaa
 4201 gcgtacttaa aggggataga tgagctctct tccttctttg ttttttttct gactggactt
 4261 ggacatgcat gtggctgggt cagccaggtt ggactgaatg gcacatgttg tggggcagag
 4321 aagaacaaga tagaagagtc tgggttcctg acaagcaggg tggcttgatc atccatgagt
 4381 tgcctttttc ccctgaacgc tgcattccac aactattgtt ggaaatcaaa atcttttttt
 4441 aagaatacta catattgctc atttctttaa tggtttcatt ctctaaggct aaaaagagtc
 4501 atggtacatg catgcatatg tgtatgtgca taagtacact ctcttttgta ctttctctac
 4561 atttaaaact gaacacagag ctccctctga atttttaacc ctcaaacac aaaaaagct
 4621 gaatatggac tttgcccag ttttatcaca ttttaataaa taaaagcttc aatgtattca
 4681 ggaattcatt atttatttgt ttcccacact ctcaacccta gattattaat ttggttcacc
 4741 taaagccaac tcacctcaaa ttggttaaaa aaaattaagc acgtgtaaat caatgtatag

FIG. 3B

4801 ggaaaggggtg gctgatccag caaaatccaa gtgtaaaaac ataaagagag gcttcacaac
 4861 agtaatgagc catctgcaat ttagaatcat ccatggtaac aatctctcat tccatttgtt
 4921 tatgttttat tcatcattgt gaaaatactt agtgaagtga caatgatgag attcaaaaat
 4981 actgcatggg tcttttttta aaaaggatat taatggcttt cagggactca actattcaaa
 5041 tgaaaatatt gctggactaa atttgctcaa atttggtcac ttcctcttca ttcttgata
 5101 taatatattc ttcattttat gacaaaaaac tatgtggtga aaacatgaga aattccgtag
 5161 tttccatcgt ttatacattt aaaaaatcaa tagcagaaac ctatttcaaa tgtacaatcc
 5221 tcattcctat catgacagag taaaccaggc tgcgggtgtc tgggtgagcg tgtgtcatt
 5281 ccctagtact acagttattt agcctcaaac gtactgtctc tagttacata tatggcaaaa
 5341 acaaggcttg acctttaaga atgaaaagaa aaaaagataa gtactgagct gttaagccaa
 5401 cttaagggta ttcattacaa agaagacaga ataactactg aacatactct ttgattctaa
 5461 tatctctatg aagaaacctc tctctacaca aattgataga cattttaaaa gaaaatttat
 5521 tcatcccaaa ttgatgcact taacttcagc tagctgattt tataatttctc aatctagctt
 5581 ataaatacta ctcttcattt acttcaattt acggagtatt tacaaaagcc ctatttcaaa
 5641 tctgggtcct atcagccccg ttttttctta tttcttctta aataagcata ctttgtaaaa
 5701 tatagccatt tattcatata tataatttctt tcatttggtt tagtaaatga catttattta
 5761 ggaaccattt tttaagaaat aagtttttac catctcaatt tgcactgaac cagcgtcttc
 5821 tgagaccatc cctctgggtg caaatctcag agtaagctca gttactccta acgactggaa
 5881 aaggaaagaa cacaggaaca cttgaactcc tctgacaagc agacactttg gctaaggcgg
 5941 cattcacatc ttctacatca tccttctctg ggaaagaagc tctcagaaaa gtccacacat
 6001 agaagagcca ctttatggtc ttccctcaat tttctcagat cagcctgatt tagtctatga
 6061 ttgtgcasac aaggagacgg agccaaggga caaaaggaag ataattgtca cacaagaaa
 6121 atgtcaaga ccttgccaag ggagagactg ctctaaataa taacaaggct attagcaaac
 6181 acaggttgag acagggaaca gaaaaagaaa agatgtcwca tgtgaatatg aaaacaaaag
 6241 ctgacaagtw attccattca gtttgccaag gcacctgcca aaaatgtccc caagggaaaa
 6301 gggccatgtg ctccatgacg ctcttcacag tgaacgttca catwaaggca cagaacattc
 6361 tgcctggcca caaaacacaa caaatatcat caatatcttg cgggggtaac agttgtgcac
 6421 tacctaccat aggcagtctc ctggtgtctc acttctcttc tcagtctgtt gtaatcttct
 6481 ttaacacttc ccagcttttt gacattcctg gaactcaaca ccaacagctt gctgaggagt
 6541 ccctccagcc cttctctctc tgatgttaat gatctaactc cctcggtgag gtctttagcc
 6601 gtccagaaat ggtttcctac agcacgttgt gggagagaag aagggaaaac aagctttaca
 6661 gcatccacac acatggaaga tctagcagca gctaattgtac attcacagca tgggcaaaga
 6721 aactgcactt tgagattatg twcaatttcc ttagcagtc tctggggaaa gaacctggaa
 6781 aaatattgct ccccttaaac tggaaaaatc accgtctttt ttatcacctt attcttgta
 6841 tgtccatttt aggagttcca ttggtgaagc catctcccgt aaagtatgat gacccaaaaa
 6901 tgtcccagtg gaatccagag aaaccacaat gtgtcactcc attccaataa tatacaaatg
 6961 tcaaagagat gacatattat agcagtgtgg aaggggactg ggccatataa tcagagttga
 7021 cctaaattaa attcctcccc tgttactcaa attgtgtgac ctggagcaag aaaattagca
 7081 catctggggc aaatttctta taaaatataa ttgcaataac acctgactca taggtagtta
 7141 caagaataaa caagaaattc tgtaaagcac tgagaaaaat atatgataaa ttaaatgtag

FIG. 3C

7201 caagattgca ggatacaagg tcaatgtaca aacatcaatc ttatgtctaa ataacaaaaa
 7261 actagaagat gatttttttaa aaaatacaat ttacaaaata ttttaaactc cataatactg
 7321 agggatactt tcttttaaata tgtgccagag ctctacagtg aaaactatct ttccaaaact
 7381 gctgaaacaa attaaagaca aataatttaa aagatacact gtatacatat actggaagtt
 7441 caatattggg aagattttcaa ttttccccac attgattcat tgatttaata taattccaat
 7501 caaaatgtta tgagggttctt ttgtagaaat ttatgtgctt attctaaaac gtctaaaagt
 7561 gagctgggca tgggtggctc acacctgtaa tcccagcatt ttgggaggcc gaggtgggtg
 7621 gtcccttgag gtcagggtgt caagaccagc ctggccaaca tggcgaaacc ctatctttac
 7681 taaaaatata aaaattaatc tggcacagtg gtaagcacct gtaatcccag ctactcaaga
 7741 ggctgaggca ggagaatcac ttaaaccagc gaggtgtagg ttgcagcgag ctgagatcct
 7801 gccactgcac tccagcttga gcaacagagt gagacttcat atctaaacaa ataaaatcta
 7861 gaagaatgag cccatgagtt tgaggctgta gtgcactatg attgcacttt tgaacaggca
 7921 ctacactctt gtctgggcaa catagcaagt cccacactca aagaaggaga gggagaaaga
 7981 agaagaagga ggaggaggag gtgaaagagg gagggaggga aggaaggaga agacgaagaa
 8041 gaatgaagaa gaataggaag aaagaagaag aaggaagagg aggaggaggga tgacccaaam
 8101 caatccttaa aaagtgagtc ttggtttcaa gactaatgct aaagccacta taaatcaagg
 8161 taatatggca ttgaatatac atagacaaac agatcgatgg aatagaataa acagtcaggg
 8221 agtggattaa cacatataca gccaaccaaa gtgccaaagc aagtcaatga ggaaaagaaa
 8281 gtattttcaa taaatggtgc tggacaact agctatctac atggaaaaaa tattaacctt
 8341 gattcatacc tcaaacacaa caaaaattta attaaagggt gatcacaggc ctacatgtaa
 8401 aagttaaaac tataaagctt ctacaaaaaa aaaatgaata tttgattgac tttggagtag
 8461 gcaagaactt cagataagat acagaaaaca ctatttttaa aaggataaat gggctttatc
 8521 aaaattaaaa acttctcatc aaaagacaca actaaaaaat gagacgacaa gtcacagact
 8581 tggagaaaat cattgcatta tgtttaactg acaagggaatt gaatactaca acttgacaat
 8641 taaaacacaa atggtgcatc tgaaaaatgg tcaaatgact taacaggcat ttacaaaagt
 8701 agacgcacaa atggctaata agcatatgaa aaaacatttg acatcattaa tcattaagaa
 8761 aagaaaaatt aaaaccataa atattccact acaagcccac tcaaatgttt caaattaact
 8821 caactaaaaa tgtcaacatt cacaagggtg tggagcaact ggaactctca cgcagtgtaa
 8881 atggtccctt caagaaaagt tctgccagtt tctcataaag ttcaatgtat acctaccctt
 8941 taatccaaca attctacttc caggcattta cacaaaagaa atgaaaacat gtcataaaaa
 9001 atatttgcac tcaataaaat ggtcacagtg tcttatgcct gagagccagc cccatactgg
 9061 aagcaacctt aatatccatc agcaggaaga tgcaaaaata atactatagc tatgcaatga
 9121 aagactactc agcaataaaa agaaacaaac tcgtcataca tgcagcaaca tgtatgtgtc
 9181 tttttctttt tttctttttt tttttttttg agacagagtt ttgttcttgt tgcccaggct
 9241 ggagtgcagt ggtacgatct cagctcactg caacctccac cttccagttt gaaactatcc
 9301 tcctgcctca gcctccaag tagctaggat tacaagtgcc caccaccagc cccagttaat
 9361 ttttgtattt tttagtagag ccgggggttc accatgttgg ccagactgtt ctcaaactcc
 9421 tgacctygtg atccaccac ctcagccttc caaagtgtgt ggattacagg cgtgagccac
 9481 cgcgccagc caacatgtac atgtcttttag aaaagatgca aaagagtagt gccacatgat
 9541 tccattaagt tctagaacag ccaaaattaa tccatggtag aagatatcag aacagaggag

FIG. 3D

9601 ggacgatgat gactgggaca agataggggtg ttgtttcttg agtgatgaga ctgttctatc
 9661 tcttggtagg gctgttcgta acacaggtgt caaaattcaa gaactgtaaa cacttaagac
 9721 ctgtgcattc cactgaatat aaaccatacc tcaattaaaa agaggaaaaa ttaagaaatt
 9781 ttttaagcatg aacattgaac aataaataaa tatataacat taatgagtgt cccataaatt
 9841 agtttagtgt aattagaaaa cccactgtg ttttctcaga atttatagta accataatca
 9901 caataatagc tacaaacttt ctccacttta aatgtcatca caacttttagg aaacagggtta
 9961 tatcttatat tattattatt attattatta ttattattat tattatttta cattttaaca
 10021 tctctgaaac agggatgtgt cctacaatca ccatcagcca ggaagtgggc aggatgcagt
 10081 gggtataact ggattatatt cctgggctg gctgaccatc tgcaaccccc tgggtgttca
 10141 atgtacaaac tatttttagga ccatctgcgt actggctgtt gctgaaaaat cttgtgttga
 10201 tgtggatttg taagactggg aaagcaccag cattcaaaac ttgaagaaag gtgttagcag
 10261 tttggcagac aatcctacag aaaactcttt attcttgggg gagaaaaagt tgccctccat
 10321 caacttcaat gggttcacagg atgatgacgt gtggggaaac acaaacctcg acaacattga
 10381 gtccaaatgt gtttcggaag actcagaatg aagaagaaga tgccaaggaa aaccttaacc
 10441 aatttatatt cctttttaat gagtgatatg tgatttttca aaagcctaaa tacatctaaa
 10501 agagctatat cagttaagta taagataaaa attctaaacc atagaaagca ctgttactgc
 10561 ttaattgaca gcaatttttc tttctttgca atgtataaaa taatgggtgc tttttatggt
 10621 cactgaaatt catattgttg tctccttta cagctgggaa agcggaaatc tggagaacta
 10681 aaatatcttg ctgaagatca caatgcagta aagcataggg caagtttcca actcgaacat
 10741 caccacagaa cacttccttt taaacataaa ttatgttgct aactttttta ttgttactta
 10801 gcaggcagat ttcaaactga tacaaatggt gaaaatgata caaaatggtt tcccacctta
 10861 aaggcattct gaaacctagc acataacatc tgtcactgaa gtccatagtt tccaatgccca
 10921 tgattggata aagacctat tgagactcat gtgtttgtct ctactgctt aaaacactca
 10981 tgtgttttaa gcaacatgac aagcacaaaa gagaacaagt actgctcctg ctctccagga
 11041 agttaccata aaataaaggg gataagggtgc cttccaaaga aaaaccaaact actaggagag
 11101 gtcagaggag aagagagatt attccagaag gagagactga ggcagatttc aggaggggag
 11161 aatataagag tcccaaaaga agaaacagga acccagggtta gagagcaatg ggcaaggggt
 11221 aggaggtgag actaaactca caaggaagtt gcttcaaaac aatagtgcct caaatgccaa
 11281 gctgatgggt tggaccatt taggaaaaac ccacaaagtc cacaggagag cagactgcac
 11341 ttccatgagc ttcaagctgt aaacagtaac ctcaaataaa acttttcccg gatcggtaat
 11401 ttccatcctg aacaatggcc aagtcaagga agctgaggat cgagccccct accgcatgcc
 11461 acaaagaacc tcaagggacc catgctctg gggagctcgc aggaggacca gcacagagggt
 11521 gtgactgcgt tggggcaggg tggccacatg tacaatgttg taccttccaa tgttctcct
 11581 taggctgtgc ttatgaagtc ccaacatgtc ttctgggctg tgttttcaac caagagagtt
 11641 aataaaaagt gggatgggca ccaactatgc ctccagagac atcaggcaga atatgccaca
 11701 gcgtgccatg tcctcctgca aacagtcaga aagaaaagat ctctaaaagc aaattggaag
 11761 agactatata tacagagcga gtgccgttac agaagtgttg ctgcaaatac accctcta
 11821 cccttgcaac ccactgcaag ctcgagagac cagtgcctt ggtaaaacac aggagagatc
 11881 attaatataa gcctaggccc tttctttgat cctgagggat aaggcttggg aaaatgcaca
 11941 gtgataaaag tgaacatttt aaaagagcag agtgagtgtg aattcccaaa gagaccattc

FIG. 3E

12001 atatcaagac ccaagccatg agaaaaatgct acaatgccct tgattaatac tttangaggc
 12061 attttgacct cagatctaata cccatgaggc ttagaccaag accttgcttg atgcagacct
 12121 ctctgactgc tgaaattggc caggagctca gtgatgaaaa tttggaaagt caatgctgat
 12181 acctgaagaa agtcccata gagatctgtg acctgaatct tcataacttt cacaggccct
 12241 ttctaaatgc caaggaggat gtgtgtggag aagaatatgc cacctatggc atcccaatat
 12301 gcctgtgccc gcctgctctt caacaaggga atgaaacctg ggtattctca taggtttttt
 12361 aaagcacttg tccacatatt ttattccgca aataacacaa tttaaaattc agctagaatt
 12421 tgccattttc tgtcaggggt gtgcatttct tccagtcacc atgacttcct gagatagcca
 12481 aaacacaaaa tttaaattgt gtttaaatta aataaacaca atttatgctg aagctgactt
 12541 taccactctt gttccaacat tacttatcct cagcatctct gaattcaa atcagccagg
 12601 tgtgggtggc catacctgta agcccagcac tttggcaagc caagatggga aggttgcttg
 12661 aggagtggga gaccagcctg ggcaacatag caagaccccc atctctacaa aaaaattaaa
 12721 aaattagcca ggcattgggt tatgcacct tagtcctagc tacttgggag gctgaggtga
 12781 gaagatcact tgagccgagg agcttgaggt ggcagtgagt catgatcata ccactgcaat
 12841 gtagcctgag caacagagca agaccccatc ccttaaaaaa aatacctgcc ctctgtactt
 12901 catagcacct ggggcatttt gaacagttct tccctcacct gaggtcatta tttgagaatc
 12961 caatttttaga gactacaatg ctctcattac ttggattgta tcatttaata ttcccaacaa
 13021 tccaggaagg ggagtctctt aatgattgtc cttgatttcc agttaaaaaa tctagacca
 13081 gaggtaacct tgtgaaggct acagagtttc agagttctaa gatcccatgc ggagattggt
 13141 gaacttgaac ttgtttcttc tgatctcatg tcaataagtg acttagtcaa cacctgggta
 13201 gaaatcatca cttaaagtga aaagcacatg gtgaggggga tgaatcctgt gccacatgg
 13261 tccctttcac ctgcccaatc ctcaggcctc tcacctggggg taccctcac acacatggca
 13321 ctttctgaa ctcatggaca tttcagtttg taacgatcct cctaagatta tctactttca
 13381 cctatccaca tttgaactat acttctttag tggaaaaaaa gaaataaagc cagtgtgag
 13441 aatgatcaaa ctccagaaat agtagactgc tgtataattt gtcttgctg acaccagatc
 13501 tttcaaggaa ctgcaagctc aagtgcacag ttaagcaaga tcaagatcta cctttgtctt
 13561 tcagagcctt cccttgggca ggtcaaggat atctgctgtg agcctgggct ggctgaacca
 13621 aagggcaatg gaggcggcca agaagcccca gccaaactgt cccctcccca accccttct
 13681 ccctggaggt cccttcagct agaatttgcc attttctgtc agggttgtgc atttcttcca
 13741 gtcaccatga ctctctgaga tagccaaaag catttgctgt gagcagagcc ccattggctg
 13801 ctttgcacat caaacatttc agacagccat atatttcata gaatacaagg ggaagataac
 13861 actggacca cagaactttc caacagcttg agaaaaatgt gtgcctgggt actgactaaa
 13921 tatacaatcg gacaacagcc aggcataagc aaaagatgga actctgatcc acaagggtgca
 13981 gcaactttcc cagaaaacaa tacccttacc cacaaaaagc aaccccgga gctggcctgc
 14041 tctaaatcag acctgctgga agccaaattg ctctctagta actatccagg aagctaaata
 14101 ataactttct taacaattga ctcaaaatgt tcaggactgg attaatcaat gatggcttcc
 14161 ctaatttttg tccccgattc caacttagga cccaccagag aaagccaaat atgggtccct
 14221 gaccaagcac acaggatgtt cgcttctagt gagccgcctc cagcttcccc gcattgataa
 14281 cctccaatca gggcacacct gaagccttcg cctttttcca ctataaagct ctcccactcg
 14341 tctgcccacc tttgagtctc tgccaaattg caaatgatgg tgtctgactc ccttgctgta

FIG. 3F

14401 gcaagctctg cataatagct tgtgcttatt ctctttgggt aggtcttcat ttcttttcac
14461 accagatagg catcatcttt gaaaacatct ttactgggtc tcttgtctcc ctctgcagc
14521 cactgatcta aatcacagac ttttttagaa agagggaaaa tagtaaagaa agaaagcgat
14581 tgagttgatg ctcccagcca ctgtctgtac cctgaatttt gtttttctgc catatgtcag
14641 gtggcctctc ctggaccatc gtcactgttt ctctgcttc cctgcagctg cagacacagg
14701 agcactgtgt aggtctggcc ttgagcacc acccaccag tacagagcca gccctgctg
14761 ttctgaatca gtgcagtgc cacttagaac acagtgccca tgctctgtg atctgaactg
14821 ctctttccag ttccatacaa tgtaagaac ccacttaaga atgttttaac cattcagtag
14881 tttttcgtgg tatataataa aaaagtgatt taaatacaat ctattttatt ttagattata
14941 tcattagggg ttattttaca aagttacctt taaatactat tgtaaataaa attaataaca
15001 aa

FIG. 3G

ID	PAIR	FORWARD	REVERSE	SEQ ID	NOS	LOCATION	SIZE
disc01	f2/r2	CATTAGCTGCTGTAGATCTTCC	TAAGGCACAGAACATTTCTGCCTG	50	51	exon9	353
disc02	f3/r3	TTATTTCAATGTGACTGCAGTTGC	CACCTGCTCTTTTCACTGATGG	52	53	exon10	280
disc03	f1/r2	GGTATGATGAACATCACTAGATTTC	CTAAGTCATCAATCTGCTCTCTCATC	54	55	exon11	191
disc04/29	f1/r1	GAGCCTACAGCTCCAGGAAG	GAGCATGGTCCCAAGACACC	56	57	exon11/11'	265
disc05	f2/r2	GAAGCTTCCCTTTGTGTTCTGTC	CGGCAGCATCAATTTTGTGGCATC	58	59	exon12	297
disc06	f1/r2	GCAGAGGCCAGGATCACC	CATCCCGTCACCTCTCAGGC	60	61	exon13	203
disc07	f1/r1	CATGACAGCTGGTGCCACGAAG	CAATCAGCTCTCTCTTCCATATTC	62	63	exon13	242
disc08	f1/r1	CAGTGTAAACTGAGGAGTCTGC	CGTGTAAACAGGTATGACAGAGTC	64	65	exon13	296
disc09	f1/r1	CATTTGCCCTTCTGCTGCAAG	CTTTCTTCAGATGCAATCATTTGCCAC	66	67	exon13	299
disc10	f1/r1	CAACCTCCAGTGATATGCCAC	CTGCCTTACTAGCACTGCACCTG	68	69	exon13	320
disc11	f2/r3	CAGTAACACAATACAGTACTCAGGC	CTGCATCTGTACTGAGCATTTGC	70	71	exon13	292
disc12	f1/r1	CTCAGGCAGTCAATACTCAGTG	ACTAGCTGCCCTGTGTTACTGAG	72	73	exon13	303
disc13	f1/r1	GTGTAGTGCTCAGTAGGACAGC	GTACTTGACCAAGAGAGGGTACTC	74	75	exon13	297
disc14	f1/r2	CAAAAGTGCACTGAGGAAGGTGG	GTGATGTAAATCAGAGTTTGGACTGG	76	77	exon13	311
disc15	f2/r3	CTCCTATTTCATATCCCATAGATCTAG	CAGAAGTGTGTGCACAGCATGG	78	79	exon13	335
disc16	f2/r2	CACCTGGCGTTTCCAGAAGGCATC	CACCTGGAATTAACCTCAAGGATTC	80	81	exon13	336
disc17	f2/r1	CATTTTCAATGATGAACCAAAATTCCTG	GCAGCAGGAAATGAAACACATATC	82	83	exon13	376
disc18	f2/r1	CTGCAAGTTAATAACTGCCTTGAATTG	CTGCTGTATATGATGATTTGAGGATAC	84	85	exon13	307
disc19	f2/r1	CAACCTTCTCAGGACAAGCCAC	GTGACCAATTTGAAAGCCAGCATTC	86	87	exon13	314
disc20	f1/r2	GTTAGAATCTGATTTGACTGGGATG	CTTGGGAGGCTGAGGCAGTAG	88	89	exon13	295
disc20	f2/r1	GCACATACTGGAATGATGAGTTAG	ATGCCCTGTAAACCCAGCTACTTTG	90	91	exon13	334
disc21	f1/r2	CTCTACCTCCAGGTTCAAGC	GGTACCAACCCGTTACATGTTTCTGG	92	93	exon13	342
disc22	f2/r2	GCACCCGGCCAACTTTCTG	CTCTGACTGTTAGGTTCACTATTAC	94	95	exon13	306
disc23	f1/r1	CTAAGGCACAGAGCTGGTAAAATG	GATGATGTGTGGGTGAATAG	96	97	exon13	292
disc25	f2/r2	AGCCTCACTGTGAAGTCTAGGC	CTATCATCCTACATTTCAAAAGGAC	98	99	exon13	403
disc26	f1/r1	CACCTGCTTCTGATTTTCTAGCTG	GAGGACAAAACACGATGTGCTGG	100	101	exon8	240
disc27	f2/r2	CTCGCTGAGGAGAAGAAAGGAGC	GGCAGCCCTGACCTACGC	102	103	exon1	316
disc28	f2/r2	CTGTAGTGGTATTGAATTTGTTTACC	CAGAAGCACTCTCTCTGGCTC	104	105	exon7	287
disc30	f2/r1	GCCAATGCTGGAAGGAACACAGC	TGCCAGCTCTCTCCGCTC	106	107	5'/promoter	168
disc31	f1/r1	CATCCCTCCATCTTCTCCATCAC	CACGCTGTGGAGCGGGC	108	109	5'/promoter	207
disc32	f2/r2	GAGTCAGTTGCCCTAGGAGCTTGC	GTGCAGGAAGCTCCAGGATG	110	111	5'/promoter	279
disc33	f1/r2	CCTATCCCTGAACCATTCGAAGAG	CAGCAGTCCGGGTGGTTTC	112	113	5'/promoter	300
disc34	f1/r2	CTGGGCCAGTAAGATCTGCATG	GAAACAGGCTCCGGACCAAG	114	115	exon2	285
disc35	f1/r1	CCAGACAGTGTGCCCTTGACTC	GGCTGCAGCTGTTGCTACTCAAC	116	117	exon2	318
disc36	f1/r1	CTGCAGCGATGGAGCAAGGC	ACATCGGGTCTCTCGTGAG	118	119	exon2	295
disc37	f1/r2	CAGAGAGCTGAGTCCCATTTG	GCACTGGTCTCCCATTTCTCTGAG	120	121	exon2	314
disc38	f1/r1	CCAGTTCTCTGGATCCCTCAC	GAGGAAGTCAGTTGAGCCCAAGAAC	122	123	exon2	246
disc39	f2/r1	CAGTTTCTAAATGTTCTTAGTTTTCAC	GATGGAAGAAAATTTGGACATGATGAC	124	125	exon3	215
disc40	f2/r1	GTTCACTACAACCTGGAGCTAAGAG	CTATCTGGAGCTGAGAGGTAGG	126	127	exon4	308
disc41	f1/r2	CATGAGGATTTCACTTCTGCATAC	GAGTATGATTTGCACCACTGCC	128	129	exon5	291
disc42	f2/r2	CGACCTGACCCCCACTGGTG	GATGCCCCACACCGTGAAG	130	131	exon6	300
disc43	f4/r2	GTTTGGTAGTTCTGGTGCATATGGC	GCATGGAAGGAATCTTGACC	132	133	exon6	327

FIGURE 4